Band Matrices with Toeplitz Inverses*

T. N. E. Greville
Mathematics Research Center
University of Wisconsin–Madison
Madison, Wisconsin 53706

and

W. F. Trench
Department of Mathematics
Drexel University
Philadelphia, Pennsylvania 19104

Submitted by Hans Schneider

ABSTRACT

It is shown that a square band matrix $H=(h_{ij})$ with $h_{ij}=0$ for $j-i>r$ and $i-j>s$, where $r+s$ is less than the order of the matrix, has a Toeplitz inverse if and only if it has a special structure characterized by two polynomials of degrees r and s, respectively.

1. INTRODUCTION

A *Toeplitz matrix* is a square matrix in which all the elements on any stripe are equal, where we follow Thrall and Tornheim [4] in defining a *stripe* as either the main diagonal or any diagonal line of elements parallel to it. More precisely, $T=(t_{ij})_{i,j=0}^{m}$ is Toeplitz if there is a sequence $\{\phi_\nu\}_{\nu=-m}^{m}$ such that $t_{ij}=\phi_{j-i}$ for $0\leqslant i,j\leqslant m$. We shall call a square matrix $H=(h_{ij})_{i,j=0}^{m}$ a *band matrix* if there are nonnegative integers r and s less than the order of the matrix such that $h_{ij}=0$ for $j-i>r$ and for $i-j>s$. We shall call such a matrix *strictly banded* if $r+s\leqslant m$. In this paper we show that a strictly banded matrix has a Toeplitz inverse if and only if it has a special structure characterized by two polynomials of degrees r and s, respectively.

*Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

0024-3795/79/050199+11$01.75

Strictly banded matrices with Toeplitz inverses have been encountered by Trench [6] in the study of stationary time series and by Greville [2] in extending moving-weighted-average smoothing to the extremities of the data.

2. THE MAIN THEOREM

We shall prove the following:

THEOREM 1. *Let*

$$H = (h_{ij})_{i,j=0}^{m}$$

be a matrix of order $m+1$ over a field F, and suppose

$$h_{ij} = 0 \qquad \text{if} \quad j-i>r \ \text{or} \ i-j>s, \tag{2.1}$$

where

$$r \geqslant 0, \qquad s \geqslant 0, \quad \text{and} \quad r+s \leqslant m. \tag{2.2}$$

Then H is the inverse of a Toeplitz matrix if and only if

$$H_i(x) = \sum_{j=0}^{m} h_{ij}x^j = \begin{cases} x^iA(x)\displaystyle\sum_{\mu=0}^{i} b_\mu x^{-\mu}, & 0 \leqslant i \leqslant s-1, \\ x^iA(x)B(1/x), & s \leqslant i \leqslant m-r, \\ x^iB(1/x)\displaystyle\sum_{\nu=0}^{m-i} a_\nu x^\nu, & m-r+1 \leqslant i \leqslant m, \end{cases} \tag{2.3}$$

where $a_0b_0 \neq 0$,

$$A(x) = \sum_{\nu=0}^{r} a_\nu x^\nu, \qquad B(x) = \sum_{\mu=0}^{s} b_\mu x^\mu, \tag{2.4}$$

and $A(x)$ and $x^sB(1/x)$ are relatively prime.

3. PRELIMINARY OBSERVATIONS AND RESULTS

A Toeplitz matrix is clearly persymmetric[1] (i.e., symmetric about its secondary diagonal), and, by considering the adjoint, it is easy to see that the inverse of a persymmetric matrix is persymmetric. Thus a necessary condition for H to have a Toeplitz inverse is that it is persymmetric. To show that H as given by (2.3) fulfills this condition, we define $\theta_{-s},\theta_{-s+1},\dots,\theta_r$ by

$$A(x)B(1/x) = \sum_{\nu=-s}^{r} \theta_\nu x^\nu,$$

and adopt the convention that $\theta_\nu = 0$ for $\nu < -s$ and for $\nu > r$. Then it is easily verified that $h_{ij} = \theta_{j-i}$ except for those elements in the $s \times r$ submatrix in the upper left corner of H and the $r \times s$ submatrix in the lower right corner. Thus it follows that H is *quasi-Toeplitz* in that h_{ij} is a function of $j-i$ alone except in these two corner submatrices. Moreover, for $0 \leqslant i < s$ and $0 \leqslant j < r$, it can be verified that

$$h_{ij} = \sum_{\mu=0}^{i} a_{j-i+\mu} b_\mu = h_{m-j,m-i}.$$

Thus H is persymmetric.

The proof of the necessity part of Theorem 1 rests on the following lemma, which follows trivially from the last four equations of [5].

LEMMA 1 (Trench). *If $H = (h_{ij})_{i,j=0}^{m}$ is the inverse of a Toeplitz matrix and $h_{00} \neq 0$, then the elements h_{ij} $(1 \leqslant i,j \leqslant m)$ are determined in terms of h_{i0} $(0 \leqslant i \leqslant m)$ and h_{0j} $(0 \leqslant j \leqslant m)$ by the recursion formula*[2]

$$h_{ij} = h_{i-1,j-1} + \frac{1}{h_{00}}\left(h_{i0}h_{0j} - h_{m-j+1,0}h_{0,m-i+1}\right), 1 \leqslant i,j \leqslant m. \tag{3.1}$$

It is also useful for the necessity proof to note that if H satisfies (2.3) and $H_i(x)$ is the generating function of the elements of the ith row as defined

[1]The term "persymmetric" is used in this sense by Wise [7], Trench [5], Huang and Cline [3], and others. Aitken [1] uses it to mean a Hankel matrix (i.e., $t_{ij} = \phi_{i+j}$).

[2]Though this formula was known long before the publication of [3], it can also be derived from Lemma 2 below by invoking the persymmetry of both H and P as defined there.

there, then by inspection,

$$
\begin{aligned}
&H_0(x)=b_0A(x), && \\
&H_i(x)=xH_{i-1}(x)+b_iA(x), && 1\leqslant i\leqslant s,\\
&H_i(x)=xH_{i-1}(x), && s+1\leqslant i\leqslant m-r,\\
&H_i(x)=xH_{i-1}(x)-a_{m-i+1}x^{m+1}B(1/x), && m-r+1\leqslant i\leqslant m.
\end{aligned}
$$

This means that

$$
h_{ij}=\begin{cases} h_{i-1,j-1}+a_jb_i, & 1\leqslant i\leqslant s,\\ h_{i-1,j-1}, & s+1\leqslant i\leqslant m-r,\\ h_{i-1,j-1}-a_{m-i+1}b_{m-j+1}, & m-r+1\leqslant i\leqslant m, \end{cases} \tag{3.2}
$$

where $1\leqslant j\leqslant m$. Conversely, if

$$
h_{i0}=a_0b_i \quad (0\leqslant i\leqslant s),\qquad h_{0j}=b_0a_j \quad (0\leqslant j\leqslant r), \tag{3.3}
$$

$$
h_{i0}=0 \quad (i>s),\qquad h_{0j}=0 \quad (j>r), \tag{3.4}
$$

and h_{ij} $(1\leqslant i,j\leqslant m)$ are computed from (3.2), then H will be of the form (2.3).

The proof of the sufficiency part of Theorem 1 rests on the following improved version of a result of Huang and Cline [3].

LEMMA 2 (Huang and Cline). *A nonsingular persymmetric matrix* $H=(h_{ij})_{i,j=0}^{m}$ *with* $h_{00}\neq 0$, *partitioned as*

$$
H=\begin{bmatrix} h_{00} & f^T\\ g & H_m \end{bmatrix}, \tag{3.5}
$$

has a Toeplitz inverse if and only if the matrix

$$
P=H_m-h_{00}^{-1}gf^T \tag{3.6}
$$

is persymmetric.

Proof. Partition H^{-1} as

$$H^{-1} = \begin{bmatrix} t_{00} & u^T \\ v & T_m \end{bmatrix}$$

where t_{00} is a scalar. Since $HH^{-1} = I_{m+1}$, it is easy to verify that $PT_m = I_m$ under the hypotheses stated here. If H^{-1} is Toeplitz, then so is T_m, and consequently $P = T_m^{-1}$ is persymmetric. Conversely, if P is persymmetric, then $T_m = P^{-1}$ is also. Since H^{-1} is also persymmetric, H^{-1} is Toeplitz by Lemma 1 of Huang and Cline [3], which states that a square matrix is Toeplitz if and only if both the entire matrix and the submatrix obtained by deleting the first row and the first column are persymmetric. ∎

In their statement of Lemma 2, Huang and Cline assumed that H_m is nonsingular. This is unnecessary.

4. PROOF OF THEOREM 1

We begin the proof of Theorem 1 with the following lemma.

LEMMA 3. *Suppose $H = (h_{ij})_{i,j=0}^m$ is of the form (2.3), with $a_0 b_0 \neq 0$. Then H is nonsingular if and only if $A(x)$ and $x^s B(1/x)$ are relatively prime.*

Proof. We assume without loss of generality that $a_r b_s \neq 0$. For sufficiency, we will show that if $A(x)$ and $x^s B(1/x)$ are relatively prime and

$$\sum_{i=0}^{m} c_i H_i(x) \equiv 0, \tag{4.1}$$

then

$$c_i = 0, \qquad 0 \leqslant i \leqslant m; \tag{4.2}$$

this implies that the rows of H are linearly independent, and so H is nonsingular. From (2.3) and elementary manipulations, we can rewrite (4.1) as

$$A(x)P(x) + A(x)x^s B(1/x)Q(x) + x^{m-r+1}B(1/x)R(x) \equiv 0, \tag{4.3}$$

where

$$P(x) = \sum_{i=0}^{s-1} c_i \beta_i(x), \tag{4.4}$$

$$Q(x) = \sum_{i=s}^{m-r} c_i x^{i-s}, \tag{4.5}$$

and

$$R(x) = \sum_{i=0}^{r-1} c_{i+m-r+1} \alpha_i(x), \tag{4.6}$$

with

$$\beta_i(x) = \sum_{j=0}^{i} b_{i-j} x^j \tag{4.7}$$

and

$$\alpha_i(x) = \sum_{j=i}^{r-1} a_{j-i} x^j. \tag{4.8}$$

Now suppose $A(x)$ and $x^s B(1/x)$ are relatively prime. Then, since $m-r+1 > s$ by (2.2), and $A(x)$ and $x^s B(1/x)$ are not identically zero because $a_0 b_0 \neq 0$, (4.3) implies that $A(x)$ divides $R(x)$ and $x^s B(1/x)$ divides $P(x)$. Therefore $R(x) \equiv 0$ and $P(x) \equiv 0$ because $\deg P(x) < \deg x^s B(1/x)$ and $\deg R(x) < \deg A(x)$.

Since $b_0 \neq 0$, it follows from (4.7) that the polynomials $\beta_i(x)$ for $0 \leqslant i \leqslant s-1$ are linearly independent, and so (4.4) and $P(x) \equiv 0$ give $c_i = 0$ for $0 \leqslant i \leqslant s-1$. Similarly, since $a_0 \neq 0$, the polynomials $\alpha_i(x)$ for $0 \leqslant i \leqslant r-1$ are linearly independent by (4.8), and (4.6) and $R(x) \equiv 0$ give $c_i = 0$ for $m-r+1 \leqslant i \leqslant m$.

Finally, replacing $P(x)$ and $R(x)$ by zero in (4.3) gives $Q(x) \equiv 0$, and so, by (4.5), $c_i = 0$ for $s \leqslant i \leqslant m-r$, and (4.2) is established.

The converse is equivalent to the assertion that H is singular if $A(x)$ and $x^s B(1/x)$ are not relatively prime. If $A(x)$ and $x^s B(1/x)$ have a nonconstant common factor, then they have a common zero ξ in some extension field $\tilde{F}$ of F. From (2.3),

$$\sum_{j=0}^{m} h_{ij} \xi^j = 0, \qquad 0 \leqslant i \leqslant m,$$

which implies that the columns of H are linearly dependent over $\tilde{F}$, and so H is singular as a matrix over $\tilde{F}$. Since nonsingularity of a matrix is invariant under field extension, H is singular over any field containing its coefficients, and so over F. ■

Proof of Theorem 1. For necessity, we assume that (2.1) and (2.2) hold and that $H=T^{-1}$, where $T=(\phi_{j-i})_{i,j=0}^{m}$. We first show that $h_{00}\neq 0$. Since $HT=TH=I_{m+1}$, we have

$$\sum_{\nu=0}^{r} h_{0\nu}\phi_{j-\nu}=\delta_{0j}, \qquad 0\leqslant j\leqslant m \tag{4.9}$$

and

$$\sum_{\mu=0}^{s} h_{\mu 0}\phi_{j+\mu}=\delta_{0j}, \qquad -m\leqslant j\leqslant 0, \tag{4.10}$$

where δ_{0j} is a Kronecker symbol. Let p be the smallest integer such that $h_{0p}\neq 0$, and consider the quantity

$$\Lambda=\sum_{\nu=0}^{r} h_{0\nu}\sum_{\mu=0}^{s} h_{\mu 0}\phi_{p+\mu-\nu}. \tag{4.11}$$

Since $h_{0\nu}$ vanishes for $\nu<p$ and (4.10) applies for $\nu\geqslant p$, (4.11) reduces to

$$\Lambda=h_{0p}.$$

On the other hand, reversing the order of summation in (4.11) gives

$$\Lambda=\sum_{\mu=0}^{s} h_{\mu 0}\sum_{\nu=0}^{r} h_{0\nu}\phi_{p+\mu-\nu},$$

which by (4.9) reduces to h_{0p} if $p=0$, and vanishes if $p>0$. Thus there is a contradiction unless $p=0$, and consequently $h_{00}\neq 0$.

Now choose a_0 and b_0 so that $a_0 b_0=h_{00}$, and define $a_1,\dots,a_r$ and $b_1,\dots,b_s$ to satisfy (3.3). By substituting (3.3) and (3.4) into (3.1), it is easy to verify that the latter reduces in this case to (3.2). Thus, the elements of H are determined by $a_0,a_1,\dots,a_r$ and $b_0,b_1,\dots,b_s$ in the same way as are the elements of a matrix of the form (2.3). Consequently, H is of the form (2.3), with $A(x)$ and $B(x)$ as in (2.4). Since H is nonsingular, $A(x)$ and $x^sB(1/x)$ are relatively prime, by Lemma 3. This proves necessity.

For sufficiency, let H be defined by (2.3) and (2.4) with $a_0b_0\neq 0$ and $A(x)$ and $x^sB(1/x)$ relatively prime, and let (2.1) and (2.2) hold. Then H is persymmetric and, by Lemma 3, nonsingular. Let $P=(p_{ij})_{i,j=1}^m$ be the matrix in (3.6), and note that the numbering of the rows and columns starts with one rather than zero. In this case f and g in (3.5) are given by

$$f^T = (b_0a_1,\dots,b_0a_r,0,\dots,0) \quad \text{and} \quad g^T = (a_0b_1,\dots,a_0b_s,0,\dots,0),$$

so

$$p_{ij} = h_{ij} - b_ia_j = h_{i-1,j-1}, \qquad 1 \leqslant i \leqslant s, \quad 1 \leqslant j \leqslant r \tag{4.12}$$

[see (3.2)], and

$$p_{ij} = h_{ij} \qquad \text{if} \quad i > s \ \text{ or } \ j > r. \tag{4.13}$$

From (4.12) and (4.13) and the fact that H is persymmetric and quasi-Toeplitz it follows that P is persymmetric. Therefore H^{-1} is Toeplitz, by Lemma 2. ∎

Incidentally, (4.12) and (4.13) show that P is the analog of H with the same polynomials $A(x)$ and $B(x)$, but with m decreased by one.

5. COMPUTATION OF H^{-1}

We close by showing how to find H^{-1} if H satisfies (2.3), where $a_0b_0\neq 0$ and $A(x)$ and $x^sB(1/x)$ are relatively prime, so that $H^{-1}=T=(\phi_{j-i})_{i,j=0}^m$ is a Toeplitz matrix. If $r=s=0$, then H is diagonal and the inversion is trivial. If $s>0$ and $r=0$, then H and H^{-1} are lower triangular, so $\phi_j=0$ if $j>0$, and by looking at the first column of $TH=I_{m+1}$, we see that

$$\phi_0 = (a_0b_0)^{-1}$$

and

$$\phi_{-j} = -b_0^{-1}\sum_{\mu=1}^{s} b_\mu\phi_{-j+\mu}, \qquad j \geqslant 1.$$

A similar argument disposes of the case where $r>0$ and $s=0$. Now suppose $r\geqslant 1$, $s\geqslant 1$, and $a_r b_s\neq 0$. By looking at the first row of $HT=I_{m+1}$ and the first column of $TH=I_{m+1}$, we see that

$$\sum_{\nu=0}^{r} a_\nu \phi_{j-\nu} = b_0^{-1}\delta_{j0}, \qquad 0\leqslant j\leqslant m, \tag{5.1}$$

and

$$\sum_{\mu=0}^{s} b_\mu \phi_{-j+\mu} = a_0^{-1}\delta_{j0}, \qquad 0\leqslant j\leqslant m. \tag{5.2}$$

In particular, (5.1) and (5.2) imply that the vector

$$\Phi = \left[\phi_{s-1},\phi_{s-2},\dots,\phi_{-r}\right]^T$$

satisfies the system

$$\begin{aligned} &\sum_{\nu=0}^{r} a_\nu \phi_{j-\nu} = b_0^{-1}\delta_{j0}, \qquad 0\leqslant j\leqslant s-1,\\ &\sum_{\mu=0}^{s} b_\mu \phi_{-j+\mu} = 0, \qquad 1\leqslant j\leqslant r. \end{aligned} \tag{5.3}$$

Therefore, if this system has only one solution, we can obtain Φ by solving it, and then compute the remaining elements of $\phi_m,\phi_{m-1},\dots,\phi_{-m}$ from (5.1) and (5.2); thus

$$\phi_j = -a_0^{-1}\sum_{\nu=1}^{r} a_\nu \phi_{j-\nu}, \qquad s\leqslant j\leqslant m,$$

and

$$\phi_{-j} = -b_0^{-1}\sum_{\mu=1}^{s} b_\mu \phi_{-j+\mu}, \qquad r<j\leqslant m.$$

If $K=(k_{ij})_{i,j=1}^{r+s}$ denotes the matrix of coefficients of the system (5.3), and

$$K_i(x) = \sum_{j=1}^{r+s} k_{ij}x^{j-1}$$

is the generating function of the elements of the ith row, then

$$K_i(x) = \begin{cases} x^{s-i}A(x), & 1 \leqslant i \leqslant s, \\ x^{i-1}B(1/x), & s < i \leqslant r+s. \end{cases} \tag{5.4}$$

We shall show that K is nonsingular, which implies that (5.3) has a unique solution. If K were singular, then some nontrivial linear combination of its rows would equal the zero vector; thus, from (5.4) there would be constants $p_0, p_1, \dots, p_{s-1}$ and $q_0, q_1, \dots, q_{r-1}$, not all zero, such that

$$A(x) \sum_{\nu=0}^{s-1} p_\nu x^\nu + x^s B(1/x) \sum_{\mu=0}^{r-1} q_\mu x^\mu \equiv 0. \tag{5.5}$$

But $A(x)$ and $x^s B(1/x)$ are relatively prime, so (5.5) implies that $A(x)$ divides $\sum_{\mu=0}^{r-1} q_\mu x^\mu$. Hence $q_0 = q_1 = \cdots = q_{r-1} = 0$, since $\deg A(x) = r$. This and (5.5) imply that $p_0 = p_1 = \cdots = p_{s-1} = 0$, a contradiction. Hence (5.3) has a unique solution.

A similar argument shows that, alternatively,

$$\Phi' = \left[\phi_s, \phi_{s-1}, \dots, \phi_{-r+1}\right]^T$$

can be found by solving the system obtained by replacing the limits on j in (5.3) with $1 \leqslant j \leqslant s$ and $0 \leqslant j \leqslant r-1$, respectively.

It is now clear that the elements of H^{-1} do not depend on m, in that with $A(x)$ and $B(x)$ given, increasing m merely enlarges the sequence $\{\phi_\nu\}$ without changing the elements already determined. Thus, corresponding to every pair of polynomials $A(x)$ and $B(x)$ of degree r and s, respectively, with $a_0 b_0 \neq 0$, such that $A(x)$ and $x^s B(1/x)$ are relatively prime, there is an infinite family of band matrices of the form (2.3) of all orders greater than or equal to $r+s$, all having Toeplitz inverses with elements taken from the sequence $\{\phi_\nu\}_{\nu=-\infty}^{\infty}$ that is the unique solution of (5.1) and (5.2).

We are grateful to the referee for his careful reading of the paper, which has resulted in some improvements in clarity.

REFERENCES

1 A. C. Aitken, *Determinants and Matrices*, 8th ed., Oliver and Boyd, Edinburgh, 1954.

2 T. N. E. Greville, Moving-weighted-average smoothing extended to the extremi-

ties of the data, MRC Technical Summary Report #1786, Mathematics Research Center, University of Wisconsin—Madison, Aug. 1977.

3 N. M. Huang and R. E. Cline, Inversion of persymmetric matrices having Toeplitz inverses, *J. Assoc. Comput. Mach.* 19:437–444 (1972).

4 R. M. Thrall and L. Tornheim, *Vector Spaces and Matrices*, Wiley, New York, 1957.

5 W. F. Trench, An algorithm for the inversion of finite Toeplitz matrices, *J. Soc. Indust. Appl. Math.* 12:515–522 (1964).

6 W. F. Trench, Weighting coefficients for the prediction of stationary time series from the finite past, *SIAM J. Appl. Math.* 15:1502–1510 (1967).

7 J. Wise, The autocorrelation function and the spectral density function, *Biometrika* 42:151–159 (1955).

Received 21 August 1978